Tom Bossuyt (Ghent)

Lice in the fur of our language? German irrelevance particles between Dutch and English

Abstract: The present paper compares the distribution of English *-ever*, German *immer* and/or *auch*, and Dutch (*dan*) *ook* in universal concessive-conditional and nonspecific free relative subordinate clauses (e.g. G. *Was auch immer du willst* 'Whatever you want') and in their elliptically reduced versions (e.g. D. *... of wat dan ook* '... or whatever'). By combining large language-specific corpora such as the *DeReKo*, *SoNaR*, and BYU corpora with the smaller multilingual *ConverGENTiecorpus*, 38,748 instances were obtained while maintaining comparability. Whereas present-day English has only one option in both clausal and elliptical constructions, viz. *WH-ever*, Dutch and German show more variation: in Dutch, discontinuous *W ... ook* is by far the most frequent option in subordinate clauses, while the complex particle *dan ook* is largely confined to elliptical constructions. In German subordinate clauses, *immer* in adjacency to the *W*-word is the most frequent option, thus corresponding to English *WH-ever*, but in elliptical constructions *auch immer* is predominates, thus corresponding to Dutch *dan ook*.

Zusammenfassung: Der vorliegende Beitrag vergleicht die Distribution von engl. *-ever*, dt. *immer* und/oder *auch* und ndl. (*dan*) *ook* in universalen Irrelevanzkonditionalen und verallgemeinernden Relativsätzen (z.B. engl. *Whatever you want* 'Was auch immer du willst') sowie in ihren elliptisch reduzierten Varianten (z.B. ndl. *... of wat dan ook* '... oder was auch immer'). Dank der Kombination großer sprachspezifischer Korpora wie des *DeReKo*, des *SoNaR*-Corpus und der BYU-Corpora mit dem kleineren mehrsprachigen *ConverGENTiecorpus* konnten 38.748 Belege erhoben werden, wobei Vergleichbarkeit gewahrt blieb. Während im heutigen Englisch *WH-ever* sowohl in Nebensätzen als auch in elliptischen Konstruktionen die einzige Möglichkeit ist, zeigen das Niederländische und Deutsche mehr Variation: In ndl. Nebensätzen kommt das diskontinuierliche *W ... ook* am häufigsten vor, während sich die komplexe Partikel *dan ook* größtenteils auf elliptische Konstruktionen beschränkt. In dt. Nebensätzen ist *W*-adjazentes, dem engl. *WH-ever* entsprechendes *immer* die häufigste Möglichkeit, in elliptischen Konstruktionen dominiert aber das dem ndl. *dan ook* entsprechende *auch immer*.

https://doi.org/10.1515/9783110668476-004

1 Introduction

Reiners (1949: 283) referred to German modal particles as "die Läuse im Pelz unserer Sprache" (lit. 'the lice in the fur of our language'), dismissing these small words, such as e.g. *mal* and *doch*, as superfluous and not worthy of the attention of linguists (cf. Hentschel 2012: 124f.). In the meantime, the tide has turned and the amount of work published on modal particles has been overwhelming ever since (König 2010: 79; Müller 2017: 384). Other kinds of apparent "lice in the fur", on the other hand, seem to have been mostly ignored. Among them are so-called "irrelevance particles", i.e. quantificational particles which occur in universal concessive conditionals (henceforth: UCCs) in certain languages (Haspelmath/König 1998):

(1) a. English: What*ever* he says, nobody listens to him.
b. German: Was *immer* er *auch* sagt, jeder hört ihm zu.
'Whatever he says, everybody listens to him.'
c. Dutch: Wat Jan *ook* zegt, Marie luistert naar hem.
'Whatever John says, Mary listens to him.'

Like all concessive conditionals, UCCs express a basic conditional meaning (König 1986; Leuschner 2006; Breindl 2014). While prototypical conditionals express one antecedent value p in their protasis, which is followed by a consequent q in the apodosis ...

(2) If the weather is nice today (= p), we'll go hiking (= q).

... concessive conditionals express a multiplicity of antecedent values (if p_x then q), whose individual truth values are irrelevant to the truth value of the consequent:

(3) Whatever tomorrow's weather is like (= p_x), we'll go hiking (= q).
a. If tomorrow's weather is A (= p_1), we'll go hiking (= q).
b. If tomorrow's weather is B (= p_2), we'll go hiking (= q).
c. ...
d. If tomorrow's weather is X (= p_x), we'll go hiking (= q).

Since these values are ordered along a given parameter – e.g. the characteristics of tomorrow's weather in (3) –, the protasis typically contains at least one contextually extreme condition p_n which carries a presupposition to the effect that $\neg q$ rather than q would normally be expected to be true (König 1986: 234). E.g. under the condition that *If there's a blizzard tomorrow* (= p_n), one would normally expect

we won't go hiking (= $\neg q$) to be true.[1] This is why UCCs such as (3) often evoke a concessive interpretation, whence the epithet "concessive".

Despite the label "*universal* concessive conditionals", the type of quantifier used in UCCs is different from standard universal quantification. Instead, it is more reminiscent of a "free-choice quantifier" (König/Eisenberg 1984: 315), whose effect is to allow the recipient to select a random value for the variable expressed by the *WH*-word in the protasis (König 1986: 231). In English at least, *whatever* and the other *-ever*-compounds are thus quantificationally more similar to free-choice *any* than to standard universal *every* or *all*, and it is precisely the "domain-widening" effect of *any* (Kadmon/Landmann 1993) that *-ever* contributes to the meaning of UCCs. With regard to German, the free-choice analysis seems to be contradicted by the fact that *immer*, usually seen as the equivalent of *always* rather than *ever* (e.g. *Er kommt immer zu spät* 'He is always late'), is used as the counterpart of *-ever* in German UCCs (cf. 1a and 1b above). There is an elegant diachronic solution to this apparent riddle, however: *immer*, which is a partial cognate of *ever*, used to have both universal and free-choice temporal readings in earlier stages of German (i.e. both 'at all times' and 'at any time'; Leuschner 1996). Its present-day use in UCCs is a residue of this earlier free-choice reading, but *immer* lost its temporal force when it was recruited as a quantificational particle in UCCs, retaining only the free-choice part of its semantics in this function (cf. ibid.: 481).

Despite the etymological link between *-ever* and *immer*, the surface realization of UCC quantification in both languages is quite different overall. As can be seen in (1b) above, *immer* is not the only irrelevance particle in German. The other option, *auch* (lit. 'also'), is etymologically identical to the Dutch irrelevance particle *ook*. Moreover, *-ever* and *immer* seem to share their preference to occupy the position immediately adjacent to the *WH-* resp. *W*-word, whereas *auch* and *ook* seem to preferably occur further to the right in the subordinate clause. We are thus faced with a rather atypical "Germanic Sandwich"-pattern (cf. van Haeringen 1956) in which German is situated between English and Dutch rather than Dutch between English and German.

Since these differences and similarities have so far mostly gone unnoticed in the literature, it is the goal of the present paper to present the first contrastive corpus study on the distribution of the irrelevance particles *-ever*, *immer* and/or

1 This extreme condition p_n is the end point of a contextually salient scale which may be either canonical or inverted. A canonical scale is invoked by (1a) above, which can be read as 'no matter how high the quality/amount of what he says, nobody listens to him'. By contrast, (1b) invokes an inverted scale: 'no matter how low the quality/amount of what he says, everybody (still) listens to him'. I am grateful to two anonymous reviewers for pointing out this and other issues.

auch, and (*dan*) *ook* in UCCs and related constructions (cf. below). This study is thus a trilingual extension of a previous study by Bossuyt/De Cuypere/Leuschner (2018) on the patterns and frequencies of German *immer/auch*, which is based on a sample of 23,299 instances with *was* 'what' and *wer* 'who' (incl. inflectional forms), gleaned from the *Deutsches Referenzkorpus* (henceforth: *DeReKo*). The latter study is in itself a semi-replication of Leuschner's study (2000) on *immer/auch*, based on 104 instances from the *Mannheimer Korpus*, which contained ca. 2.2 million tokens.

In order to obtain a large amount of sufficiently comparable data, large language-specific corpora were triangulated for the present study with the respective components of a small but comparable multilingual corpus (cf. section 2). All occurrences were analyzed for combinatorial variation on the one hand (answering the question which particle(s) is/are used) and positional variation on the other hand (which position(s) the particle(s) tend to occupy). Focusing on German and Dutch, section 3 presents the distributional patterns of irrelevance particles in subordinate clauses (3.1) and elliptically reduced constructions (3.2). A discussion on the similarities and differences between, first, Dutch (*dan*) *ook* and German *auch* (*immer*), and then between English *-ever* and German *immer* follows in section 4.

This paper will argue that the synchronic distributions of the particles represent a snapshot of the long-term emergence of irrelevance marking as a subsystem in each of the three languages, with varying degrees of grammaticalization. Whereas the grammaticalization of English *WH-ever* is more or less complete, the grammaticalization process of German *W immer/auch* subordinators seems to have lost its former directionality, resulting in a situation resembling a long-term "grammaticalization building-site" (*Grammatikalisierungsbaustelle*, Leuschner 2006; cf. Nübling 2005). Finally, discontinuous Dutch *W … ook* shows only weak signs of grammaticalization.

2 Methodology: corpus triangulation

2.1 Corpora and search queries

As mentioned above, this study combines data from very large, language-specific corpora (which are, however, barely comparable) with data from a small, but very comparable multilingual corpus. The goal of this methodology is to obtain a large amount of data while maintaining comparability. The following language-specific corpora were used:

- The "Archiv W" of the *DeReKo* is the main reference corpus of contemporary written German, containing approximately 9.2 billion tokens in total as of August 2018.[2] The corpus consists of a variety of text types, often from printed news media in Germany, Austria, and the German-speaking part of Switzerland, recently supplemented with a considerable amount of Wikipedia articles and discussions as well as parliamentary minutes (Kupietz/Lüngen 2014).
- The *SoNaR* corpus is a 500-million-word reference corpus of contemporary written Dutch.[3] It consists of both conventional media (e.g. newspapers) and new media (e.g. tweets, blogs, or chat conversations), and is fairly well-balanced between Dutch and Flemish texts (Oostdijk et al. 2013).
- The BYU corpora are probably the most widely used online corpora for English.[4] This study combines data from the *BYU-BNC* (100 million tokens of British English, 1980s-1993), *COCA* (560 million tokens of American English, 1990–2017), *Strathy Corpus* (50 million tokens of Canadian English, 1970s–2000s), *Wikipedia Corpus* (1.9 billion tokens, 2012–13), and *Hansard Corpus* (1.6 billion tokens of British parliamentary minutes, 1803–2005), containing over 4.2 billion tokens in total. Combining these corpora somewhat mimics the composition of the *DeReKo*.

The small but comparable multilingual corpus used for the present study is the *ConverGENTiecorpus*, which is institutionally available at Ghent University.[5] It consists of seven subcorpora in English, Dutch, German, French, Spanish, Italian, and Portuguese, containing about 1.5 million tokens each. Comparability is guaranteed, as all subcomponents contain approximately the same amount of tokens distributed over a wide variety of corresponding text genres.

Search queries for the present study in the *ConverGENTiecorpus* included virtually all *WH*-words, as was the case in Leuschner's original study, which, however, referred exclusively to the German *Mannheimer Korpus* (Leuschner 2000). By contrast, the search queries in the large corpora were limited to *WH*-words for 'what' and 'who' (incl. inflectional forms, if applicable, e.g. *whom*) for practical reasons, as was also the case in Bossuyt/De Cuypere/Leuschner (2018).

2 https://www.ids-mannheim.de/cosmas2/projekt/referenz/archive.html (last accessed: 19-3-2019).
3 https://portal.clarin.nl/node/4195 (last accessed: 19-3-2019).
4 https://corpus.byu.edu/ (last accessed: 19-3-2019).
5 http://research.flw.ugent.be/en/projects/convergentiecorpus (last accessed: 19-3-2019).

For the English data, search queries for *whatever*, *whoever*, *whomever*, and *whosever*[6] were conducted separately in each of the abovementioned corpora. A total of 4,642 instances were found. Search queries for these and other *-ever*-compounds conducted in the *ConverGENTiecorpus* resulted in a total of 1,240 exported instances.

For the German data, Leuschner's (2000) conclusions on the positional tendencies of *immer* and *auch* were taken into account when designing the corpus search queries, in order to maximize recall ratios. For *immer*, only instances where the *W*-word (i.e. *was*, *wer*, *wem*, or *wen*) immediately precedes *immer* were initially included. In a later stage, search strings with *immer* immediately preceded by a 3rd person singular pronoun which was in turn immediately preceded by the *W*-word (e.g. *wer es immer*) were included to find rare occurrences in which *immer* follows the subject rather than preceding it. For *wessen*, which can modify nouns (e.g. *wessen Haus* 'whose house'), the distance operator was set to 3. For *auch*, a distance operator of 4 was found to be the best balance between precision and recall (cf. Bossuyt/De Cuypere/Leuschner 2018: 101 fn. 5). A total of 53,732 instances were exported and analyzed manually (cf. 2.2 below). In the *ConverGENTiecorpus*, distance operators allowing up to 5 words between the *W*-word and the irrelevance particle were included, guaranteeing that virtually every instance in the corpus was included in the sample.

For the Dutch data, the queries were designed to resemble those used to search instances of *auch* in the *DeReKo*. This means that search queries allowed up to three words between the *W*-word (i.e. *wat*, *wie* or *wiens*) and the irrelevance particle *ook*. A total of 30,895 instances were exported and analyzed manually (cf. 2.2 below). As with the German data, distance operators allowing up to 5 words were included in the search queries in the *ConverGENTiecorpus*, assuring that nearly every instance was included in the sample.

2.2 Manual analysis of the German and Dutch data

Whereas the results for English *WH-ever* are mostly unambiguous, the German and Dutch data needed manual analysis to check whether *immer*, *auch* resp. *ook* did indeed function as irrelevance particles. This is because *immer* can also be

6 Unfortunately, the possessive form *whoever's* could not be included in this study, since this search query resulted in too many invalid instances consisting of *whoever* followed by the contracted form of *is*. Adding the noun tag did not solve this problem, nor did tagging *whoever's* as a possessive determiner.

a temporal adverb (cf. above), as shown in (4), and *auch*/*ook* can also be focus particles, as shown in (5) and (6):

(4) #Was *immer* bleiben wird, ist mein Code civil. (Die Zeit (Online-Ausgabe), 25-2-2010)
'What will always remain, is my Code civil.'

(5) #Was es heute jedoch *auch* häufiger gibt, sind Mütter, die arbeiten. (Braunschweiger Zeitung, 12-9-2008)
'However, what is nowadays more common as well are working mothers.'

(6) #Wat *ook* speciaal zal zijn, is het Japanse theehuisje van S. D. (WR-P-P-G-0000666221)
'What will be special, too, is S. D.'s Japanese tea cottage.'

Moreover, numerous doubles had to be removed from the *DeReKo* and *SoNaR* data. This brought the final *DeReKo* sample to 23,299 instances (also used in Bossuyt/De Cuypere/Leuschner 2018) and the final *SoNaR* sample to 9,305 instances. The *ConverGENTiecorpus* contains 91 instances for German and 171 for Dutch.

Not all of these instances represent prototypical UCCs as mentioned in (1) and (3). The German sample in particular contains a considerable amount of non-specific free relatives (henceforth: NFRs), as in (7):

(7) Wer *immer* bisher als "künftiger Papst" ins Konklave ging, kam als Kardinal wieder heraus. (Nürnberger Nachrichten, 14-10-2003)
'Whoever entered the conclave as a "future pope" so far, came out again as a (mere) cardinal.'

The free-choice semantics and quantificational strategies in these subordinate clauses are the roughly same as in UCCs, but the syntactic function of the subordinate clause in the complex sentence is different: whereas UCCs typically function as a loose adjunct to their apodosis, a NFR typically functions as an embedded argument in its respective main clause (Leuschner 2005), e.g. as its subject in (7), with a broad transitional zone of surface variation linking the two sentence types (Leuschner 2005: 59–62; Breindl 2014: 981f.). For the present study, however, the relevant syntactic distinctions are less important than the semantic-functional overlap between UCCs and NFRs, as shown by the fact that both clause types can be paraphrased by an open conditional (cf. Lehmann 1984: 339):

(7)' If x went into the conclave as a "future pope", x came out again as a cardinal.

It is the presence of a variable in the underlying conditional relationship that motivates the quantificational strategies that are shared by UCCs and NFRs. Details in the surface realization of irrelevance marking may well vary with the syntactic status of the subordinate clause and such potential patterns should be addressed in future research into irrelevance marking. Only the overall patterns of irrelevance marking are in the focus of the present study, however, and hence no systematic distinction will henceforth be drawn between UCCs and NFRs. This decision is reflected in the label *primary irrelevance constructions* for both clause types together (as opposed to secondary irrelevance constructions, which occur at the sub-clausal level, cf. below).

Whereas there is only one strategy to mark free-choice quantification with an irrelevance particle in English primary constructions, namely by attaching *-ever* to the *WH*-word,[7] the same quantificational effect is conveyed by different particles resp. particle combinations in different positions in German and Dutch. In order to account for this variation, Dutch and German primary constructions are analyzed using Leuschner's (2000) adaptation of the Topological Field Model (cf. Wöllstein 2014) as demonstrated in Table 1.

Table 1: Leuschner's (2000: 345) adaptation of the Topological Field Model for primary irrelevance constructions in which the *W*-word is not the subject of the subordinate clause, exemplified by (1b)

–	**pre-field**	**left bracket**		**middle-field**		**right bracket**	**post-field**
–	W	–	II	S	IV	V	–
(1b)	was	–	*immer*	er	*auch*	sagt	–

While the *W*-word occupies the pre-field, leaving the left bracket unoccupied in Standard German (Wöllstein 2014: 32–37), the middle-field is divided into a field S for the subject of the subordinate clause and two fields which may be occupied by irrelevance particles: field II to the left of S and field IV to the right of S (Leuschner

7 The *WH-so-ever*-pattern (e.g. *whosoever*, *whatsoever*) is unproductive and archaic in present-day English. The only exception is *whatsoever* as a post-nominal NPI, e.g. *no idea whatsoever* 'no idea at all'. Its intensifying meaning is, however, considerably different from the free-choice quantificational readings the present study is concerned with, and will therefore not be considered any further.

2000: 345). As usual in German subordinate clauses, the verb occupies the right bracket (V) and the post-field is standardly left unoccupied.

The topological model of Table 1 only makes sense if the *W*-word is not the subject of the subordinate clause. If the *W*-word is the subject, on the other hand, there is no need to split up the middle-field, which is then simply called II/IV (Leuschner 2000: 345f.).

Table 2: Leuschner's (2000: 346) adaptation of the Topological Field Model for primary constructions in which the *W*-word is also the subject of the subordinate clause, exemplified by (8), taken from the *SoNaR* corpus

–	pre-field	left bracket	middle-field	right bracket	post-field
–	W	–	II/IV	V	–
(8)	wie	–	morgen *ook*	wint	–

While these two models fit the majority of the data, a considerable amount of instances containing irrelevance marking does not fit either model (93/171 = 54.39% in the Dutch component of the *ConverGENTiecorpus*, 3,921/9,305 = 42.14% of *SoNaR* data; 35/91 = 38.46% in the German component of the *ConverGENTiecorpus*, 4,926/23,299 = 21.14% of *DeReKo* data). These instances are derived historically from primary constructions, but have been reduced by ellipsis (Breindl 2014: 980f.; Leuschner 2013: 57; Waßner 2006: 386f.). They are labeled *secondary irrelevance constructions* in the present study and may function as:

(9) general extenders (Overstreet 1999: 122–124, 147; Brinton 2017: 273–278)
   a. Zij worden nooit voor dief of *wat dan ook* uitgescholden.
   (WR-P-P-G-0000427484)
   'They are never called thieves or whatever.'
   b. Ich war immer betrunken, stoned oder *was auch immer*.
   (Braunschweiger Zeitung, 1-7-2010)
   'I was always drunk, stoned, or whatever'.
   c. [...] there may be a gunboat, or *whatever* – I do not know. (Hansard90)

(10) discourse markers (Brinton 2017: 268–282 on English *whatever*)
   a. [...] maar *wat dan ook*, jij bent de mooiste. (WR-U-E-A-0000104003)
   'but whatever, you are the most beautiful.'
   b. Doch *was auch immer*: Ein Crash ist trotzdem jederzeit möglich.
   (Die Südostschweiz, 22-10-2006)
   'But whatever: a crash is nevertheless a possibility at all times.'

c. [...] we'd just talk about, I don't know [pause] *whatever*, she'd probably agree with everything I said as well because that's what Catherine's like (BNC KP4 S_conv)

(11) indefinite pronouns (cf. Haspelmath 1997: 139, 160f.)

a. De beklimming van de Everest is voor *wie dan ook* superzwaar. (WS-U-E-A-0000000442)
'Climbing the Everest is super tough for anyone (lit. whoever).'

b. Ein Appell an *wen auch immer*, der sich verantwortlich fühlt. (Süddeutsche Zeitung, 17-7-2008)
'A call to anyone (lit. whoever) who feels responsible.'

c. Romney can run a great campaign, spend untold millions in the final days, do *whatever*, but it's still the president who has more agency here. (EN_ Jou_Com_0077)

Indefinite pronouns of the type *WH* + particle(s) in (11) are more common in Dutch (cf. Hoeksema 2012 on *W dan ook*-pronouns); some native speakers of German and English might even not accept (11b) resp. (11c) as grammatical. In English, indefinite pronouns from the *any*-series are usually used in these contexts (cf. Haspelmath 1997).

Since irrelevance particles show strikingly different distributional patterns in primary and secondary irrelevance constructions in German and Dutch, a clear distinction between primary and secondary irrelevance constructions will be made in the following sections.

3 Distributional patterns

3.1 Primary irrelevance constructions

3.1.1 Dutch

Table 3 represents the distribution of the Dutch irrelevance particle *ook* in primary irrelevance constructions in the *ConverGENTiecorpus*.[8] An example of each type from the corpus is given in (12).[9]

8 Note that the left bracket and the post-field are left out of this and subsequent tables, as they are irrelevant to the particles' distribution.

9 Cases with a copula as in (12a) and (12b) were analyzed as *W II S IV V*-patterns, since the finite verb agrees in number with the NP, not the *W*-word: *Wie hij.SG ook mocht.SG zijn* 'Whoever he might have been', but *Wie zij.PL ook mochten.PL zijn* 'Whoever they might have been'.

Table 3: Distribution of irrelevance particles in Dutch primary constructions with W ≠ S in the *ConverGENTiecorpus*. # stands for raw frequencies, % for relative frequencies

–	W	II	S	IV	V	#	%
a.	W	*ook*	S	–	V	3	3.90%
b.	W	–	S	*ook*	V	74	96.10%
						77	100.00%

(12) a. Wat *ook* het statuut van het kind in kwestie is, [...]: elk kind heeft recht op huisvesting, onderwijs, gezondheidszorg, ... (NE_Jou_Com_1047)
'Whatever the status of the child in question is: every child has a right to housing, education and health care.'
b. Maar wie hij *ook* mocht zijn of geweest was, hij was dood.
(NE_Jou_Com_ 1137)
'But whoever he may have been or had been, he is dead.'

Ook clearly occurs much more often in field IV (96.10%) than in field II (3.90%). This rightward tendency is confirmed by the data from the much larger *SoNaR* corpus (4,808/4,977 = 96.60% in field IV vs. 169/4,977 = 3.40% in field II), as shown in Table 4. Apart from *ook*, the much rarer particle combination *dan ook* occurs in primary irrelevance constructions and shares *ook*'s preference for field IV (132/136 = 97.06% in field IV vs. 4/136 = 2.94% in field II). (13) provides an example of each type from the corpus.

Table 4: Distribution of irrelevance particles in Dutch primary constructions with W ≠ S in the *SoNaR* corpus

–	W	II	S	IV	V	#	%
a.	W	*ook*	S	–	V	169	3.31%
b.	W	–	S	*ook*	V	4,808	94.03%
c.	W	*dan ook*	S	–	V	4	0.08%
d.	W	–	S	*dan ook*	V	132	2.58%
						5,113	100.00%

(13) a. Wat *ook* de directe oorzaak mag zijn waardoor het vredesproces is vastgelopen, het lijdt geen twijfel dat hervatting van een dialoog de spanning tot normale proporties kan terugbrengen.
(WR-P-P-I-0000000313)

'Whatever may be the direct cause which got the peace negotiations bogged down, there is no doubt that resuming the dialogue will bring the tensions back to normal proportions.'

b. De beste ploeg zal winnen. Wie dat *ook* is, ik zal altijd een fles champagne opentrekken. (WR-P-P-G-0000642881)
'The best team will win. Whoever that is, I will pop a bottle of champagne in any case.'

c. Wat *dan ook* de oorzaak is, leg de zieke met de voeten omhoog en zorg dat hij voldoende lucht krijgt. (WR-P-P-H-0000061428)
'Whatever the cause is, lay down the sick person with their feet up and make sure they get enough air.'

d. [...] om de aandacht te trekken van de geïnteresseerden, wie dat *dan ook* mogen zijn. (WR-P-P-G-0000265835)
'to draw the attention of those who are interested, whoever that may be.'

Table 5 represents the distribution of Dutch irrelevance particle(s) (*dan*) *ook*, based on the *SoNaR* corpus, in primary constructions in which the *W*-word is also the subject of the subordinate clause. (14) gives an example of each type from the corpus.[10]

Table 5: Distribution of irrelevance particles in Dutch primary constructions with W = S in the *SoNaR* corpus

–	W	II/IV	V	#	%
a.	W	*ook*	V	257	94.83%
b.	W	*dan ook*	V	14	5.17%
				271	100.00%

(14) a. Wat hier *ook* wordt besloten, ik ben ervan overtuigd dat we een onomkeerbaar proces in gang zetten waardoor heel Europa een geheel ander aanzien zal krijgen. (WR-P-P-I-0000000272)

10 There are no instances of primary constructions in which the *W*-word is the subject of the subordinate clause in the *ConverGENTiecorpus* sample, although this is partially due to the fact that instances with Dutch *er*, e.g. *wat er ook gebeurt* 'whatever happens', were classified as instances of the *W II S IV V*-pattern (cf. Table 3).

'Whatever is decided here, I am convinced that we set in motion an irreversible process which will totally alter the face of Europe.'

b. Wie *dan ook* mij deze ketting gaf, moet van mij gehouden hebben. (WR-P-E-G-0000010823)
'Whoever gave me this necklace, must have loved me.'

Despite *ook* clearly being more frequent in both types of primary irrelevance constructions, a chi-square test suggests that *dan ook* is significantly overrepresented in *W II/IV V*-constructions as shown in Table 6 (Yates $\chi^2 = 5.08$; df = 1; p = 0.02; Cramér's V = 0.33):

Table 6: Chi-square test comparing occurrences of Dutch irrelevance particles *ook* and *dan ook* in both types of primary irrelevance constructions in the *SoNaR* corpus. Standardized residuals are given in brackets, values higher than |2| indicate a significant deviation from the expected cell value and are in bold. No cells have an expected value below 5

–	ook	dan ook	total
W II S IV V	4,977 (+0.08)	136 (–0.5)	5,113
W II/IV V	257 (–0.37)	14 **(+2.17)**	271
total	5,234	150	5,384

3.1.2 German

Table 7 shows the distribution of the German irrelevance particles *immer* and *auch* and their combinations, based on the *ConverGENTiecorpus*, in primary constructions in which the *W*-word is not the subject. An example of each type is provided in (15).

Table 7: Distribution of irrelevance particles in German primary constructions with W ≠ S in the *ConverGENTiecorpus*

–	W	II	S	IV	V	#	%
a.	W	*auch immer*	S	–	V	13	26.53%
b.	W	*immer*	S	–	V	22	44.90%
c.	W	*immer*	S	*auch*	V	8	16.33%
d.	W	–	S	*auch*	V	6	12.24%
						49	100.00%

(15) a. Wie *auch immer* man Neanderthaler sehen mag. Das extrem wechselhafte Bild spiegelte immer auch den Zeitgeist der jeweiligen Epoche wieder. (GE_Sci_Pop_0464)
'However one may view Neanderthals, their extremely variable image always reflected the *Zeitgeist* of the respective period.'
b. Wann *immer* ein Land in die Krise gerät, werden seine Bürger panisch die Konten räumen. (GE_Jou_Com_0767)
'Whenever a country plunges into a crisis, its citizens will empty their bank accounts in panic.'
c. Was *immer* er *auch* jetzt sagen könnte, er müßte sich festlegen. (GE_Lit_ Fic_0005)
'Whatever he could say now, he would have to make a decision.'
d. Wo Forscher *auch* hinsehen, überall entdecken sie bisher unbekannte Arten. (GE_Sci_Pop_0630)
'Wherever scientists look, they discover previously unknown species everywhere.'

As can be seen from Table 7, the preferred position of irrelevance particles in German is clearly field II (71.43%) rather than field IV (12.24%). In fact, there are more instances where both fields are occupied (= type c; 16.33%) than instances where field IV is the only occupied field. The only particle that prefers field IV is *auch*, similarly to Dutch *ook* (cf. above).

These general distributional tendencies are confirmed in the much larger sample from the *DeReKo*, as represented in Table 8. (16) provides an example of each type from the corpus.

(16) a. Was *auch* die Gründe sein mögen, nur jammern [...] hilft auch nicht weiter. (St. Galler Tagblatt, 2-10-2001)
'Whatever the reasons may be, just complaining won't help either.'
b. Wen *auch immer* man fragt: Esel finden alle irgendwie klasse. (Süddeutsche Zeitung, 3-6-2006)
'Whoever you ask: everyone thinks donkeys are great somehow.'
c. Wer *immer auch* die Täter sind, [...], sie müssen sich vorsehen. (Die Südostschweiz, 21-4-2010)
'Whoever the perpetrators are, they have to watch out.'
d. Was *immer* sie tun, Maitressen haben einen schlechten Ruf. (Süddeutsche Zeitung, 15-4-2014)
'Whatever they do, mistresses have a bad reputation.'

e. Doch was *immer* er *auch* tut, es reicht nicht. (die tageszeitung, 19-11-2013)
'But whatever he does, it is not enough.'
f. Mit wem ich *auch* rede, überall höre ich dasselbe. (plenary minutes, Berlin, 28-6-2001)
'Whoever I talk to, I hear the same everywhere.'
g. Wessen Socke das *auch immer* ist, es wird langsam langweilig. (Wikipedia Discussion Forums, 2011)
'Whoever's sock that is, things are beginning to get boring.'
h. Zeitgemäße Dienstvereinbarungen, was das *immer auch* heißen möge. (plenary minutes, Sankt Pölten, 4-10-2001)
'Contemporary service contracts, whatever that may be.'
i. zu AC. @Hajog oder O. oder wer das *immer* ist. (Wikipedia Discussion Forums, 2011)
'to AC. @Hajog or O. or whoever that is.'

Table 8: Distribution of irrelevance particles in German primary constructions with W ≠ S in the *DeReKo*

–	W	II	S	IV	V	#	%
a.	W	auch	S	–	V	22	0.24%
b.	W	auch immer	S	–	V	954	10.53%
c.	W	immer auch	S	–	V	149	1.64%
d.	W	immer	S	–	V	6,075	67.05%
e.	W	immer	S	auch	V	1,005	11.09%
f.	W	-	S	auch	V	647	7.14%
g.	W	-	S	auch immer	V	154	1.70%
h.	W	-	S	immer auch	V	15	0.17%
i.	W	-	S	immer	V	39	0.43%
						9,060	100.00%

The types represented in the *ConverGENTiecorpus* (cf. (15a-d) above) are precisely the four most frequent ones in the *DeReKo*, viz. *immer* occupying field II (67.05% in the *DeReKo*), *immer … auch* straddling the subject field (11.09%), *auch immer* occupying field II (10.53%), and *auch* occupying field IV (7.14%). All other types, which account for less than 2% each and for only about 4.18% combined, are instances of the particles (or particle combinations) occupying their respective dispreferred field(s). Moreover, the basic tendency is confirmed that irrelevance marking in field II only (79.47%) is preferred over marking in both fields simultaneously (11.09%) or in field IV only (9.44%).

The most striking difference, however, is the proportion of *immer* in field II in the *DeReKo* (67.05%) if compared to both the *ConverGENTiecorpus* (44.90%) and Leuschner's study based on the *Mannheimer Korpus* (34/92 instances = 36.96%, Leuschner 2000: 348). A one-tailed two-proportions Z-test suggests that the proportion in the *DeReKo* deviates significantly from the corresponding proportions in the *ConverGENTiecorpus* and the *Mannheimer Korpus* ($p < 0.0001$ in both), while the *ConverGENTiecorpus* and *Mannheimer Korpus* do not deviate significantly from each other ($p = 0.18$).[11] There are several potential explanations for this difference:

1. Whereas Leuschner's (2000) sample and the *ConverGENTiecorpus* contain search results for virtually all *W*-words, the *DeReKo* sample is limited to *was* and *wer* (incl. inflectional forms; cf. above). This means that almost all *W*-words that can form complex *W*-phrases, such as *welch-* (*welches Haus* 'which house') or *wie* (*wie schön* 'how beautiful'), are excluded from the *DeReKo* sample. In fact, the only *W*-word in the *DeReKo* sample that can build complex phrases is *wessen*, which is by far the least frequent *W*-word in the sample (n = 252 or 1.08% of the total *DeReKo* sample). On the other hand, *welch-* and *wie* are the two most frequent *W*-words in the *ConverGENTiecorpus* sample, making up 49% of its instances.

   Since *immer* is only very rarely attested with complex *W*-phrases (cf. further below), but occurs very frequently with simple *W*-words such as *was* and *wer*, the difference in *W*-word coverage between the *DeReKo* on the one hand and the *ConverGENTiecorpus* and *Mannheimer Korpus* on the other hand may largely explain the proportional differences between these corpora. At a later stage of the investigation, *welch-* will be added to the *DeReKo* sample, presumably resulting in an overall lower proportion of *immer*.
2. The distance operator of 1 in *DeReKo* search queries for *W immer* may have caused *immer* to be somewhat overrepresented in this sample. Since larger distance operators make the recall ratios less precise, it is easier to find instances of *W immer* compared to e.g. *W … auch* with a distance operator of 4.
3. Tendencies relating to text genre may play a role here. The relative portion of written press texts in the *DeReKo* is much larger than in the more balanced *ConverGENTiecorpus* and in the *Mannheimer Korpus*, which contained a larger

11 The proportional difference between the *DeReKo* on the one hand and the *ConverGENTiecorpus* and *Mannheimer Korpus* on the other hand remains significant after a Bonferroni correction was carried out, which is used to counteract the increased risk of false positives when comparing more than two samples with a two-proportions Z-test. I thank Dr. Ludovic De Cuypere (Ghent) for introducing me to this method. Although Z-tests require independent data and the *Mannheimer Korpus* is included in the *DeReKo*, the enormous size difference between these two corpora (cf. above) nullifies this issue.

proportion of literary texts. To test this hypothesis, the proportions of *immer* in press texts, parliamentary minutes, and Wikipedia-texts were compared in two randomly drawn subsets containing 10% of the instances in Table 8 (n = 906) for constructions with W ≠ S and in Table 9 (n = 931, cf. below) for constructions with W = S.[12] While text genre is often a large source of unwanted noise in corpus linguistics, its role seems to be fairly minor in this case: only the difference between press texts and Wikipedia-texts in *W II/IV V*-constructions proved to be significant (two-tailed two-proportions Z-test $p < 0.0001$). While these findings suggest that the proportion of *immer* is not strikingly different across different text genres, it may still be worthwhile for further research to look into the effects of text genre on particle distribution, based on multiple genres and larger samples.

4. It is conceivable that the proportional differences between Leuschner's (2000) sample based on the *Mannheimer Korpus*, which was compiled in the 1960s, and the *DeReKo* sample, which consists mostly of texts from the 1990s–2010s, reflect a microdiachronic change. This is, however, rather unlikely, since the *ConverGENTiecorpus* consists of texts published from the 1990s until 2015, and yet shows a distribution similar to the *Mannheimer Korpus*. Another reason why the microdiachronic hypothesis is implausible, is that irrelevance particles in German are part of a larger "grammaticalization building-site" (Leuschner 2006; cf. Nübling 2005), and therefore unlikely to undergo dramatic changes within a few decades (cf. further below).

Table 9 represents particle distributions, based on the *DeReKo* data, in those primary irrelevance constructions in which the *W*-word is also the subject of the subordinate clause. An example of each type with the verb *passieren* 'to happen' is given in (17).[13]

(17) a. Denn was *auch* passiert: Freilichtspiele sind immer ein Erlebnis. (Mannheimer Morgen, 16-6-2001)
'For whatever happens: open-air shows are always a great experience.'

b. Was *auch immer* passiert, es muss schnell geschehen. (Luxemburger Tageblatt, 28-6-2011)
'Whatever happens, it has to happen fast.'

12 I am grateful to an anonymous reviewer for suggesting this method to me.

13 The *ConverGENTiecorpus* contains only 3 instances of *immer* in *W II/IV V*-constructions (42.86%), 2 with *auch immer* (28.57%), and 2 with *auch* (28.57%). Since the total number of occurrences is so low (n = 7), little can be said about these instances and they will not be discussed any further.

c. Was *immer auch* passiert, Gott will, daß wir glücklich sind. (Neue Kronen-Zeitung, 24-1-1995)
'Whatever happens, God wants us to be happy.'
d. Was *immer* passiert, wir sind bereit zu kämpfen. (St. Galler Tagblatt, 15-2-1999)
'Whatever happens, we are prepared to fight.'

Table 9: Distribution of irrelevance particles in German primary constructions with W = S in the *DeReKo*

–	W	IV	V	#	%
a.	W	*auch*	V	79	0.85%
b.	W	*auch immer*	V	1,295	13.91%
c.	W	*immer auch*	V	640	6.87%
d.	W	*immer*	V	7,299	78.37%
				9,313	100.00%

As with *W II S IV V*-constructions in Table 8, *immer* is the most frequent irrelevance particle in Table 9 (78.37%), but *auch immer* (13.91%) is more common than *immer* (...) *auch* (6.87%) in *W II/IV V*-constructions. *Auch* occurs only marginally in the latter subordinate clause type (0.85%). In accordance to these observations, a chi-square test with standardized residuals, as shown in Table 10, suggests that *immer* and *auch immer* occur significantly more often in *W II/IV V*-constructions, whereas *auch* and *immer* (...) *auch* show a strong preference for the *W II S IV V*-constructions ($\chi^2 = 735.97$; df = 3; $p < 0.0001$; Cramér's V = 0.20).

Table 10: Chi-square test comparing occurrences of *immer*, *auch*, *auch immer*, and *immer* (...) *auch* in both types of German primary irrelevance constructions in the *DeReKo*. Standardized residuals are given in brackets, no cells have an expected value below 5

–	*immer*	*auch*	*auch immer*	*immer* (...) *auch*	total
W II S IV V	6,114 **(–6.15)**	669 **(+15.63)**	1,108 **(–2.24)**	1,169 **(+9.27)**	9,060
W II/IV V	7,299 **(+6.07)**	79 **(–15.41)**	1,295 **(+2.2)**	640 **(–9.15)**	9,313
total	13,413	748	2,403	1,809	18,373

3.2 Secondary irrelevance constructions

3.2.1 Dutch

Table 11 represents the distribution of the Dutch irrelevance particle(s) (*dan*) *ook* in secondary constructions in the *ConverGENTiecorpus*. Particle distributions in the *SoNaR* corpus are given in Table 12. Examples from the corpora are provided in (18) resp. (19).

Table 11: Distribution of irrelevance particles in Dutch secondary constructions in the *ConverGENTiecorpus*

–	***ook***	***dan ook***	**total**
#	38	55	93
%	40.86%	59.14%	100.00%

(18) a. Als een rode draad door zijn politiek loopt tenslotte zijn constante weigering om welk akkoord *ook* te sluiten (NE_Jou_New_0715)
'A central theme in his politics is after all his permanent refusal of signing any agreement (lit. whichever agreement)'
b. Alleen het lezen van deze letters in het Frans of welke andere taal *dan ook* leidt tot verbazingwekkende resultaten. (NE_Cor_Pro_0016)
'Simply reading these letters in French or in whichever other language leads to amazing results.'

Table 12: Distribution of irrelevance particles in Dutch secondary constructions in the *SoNaR* corpus

–	***ook***	***dan ook***	**total**
#	975	2,946	3,921
%	24.87%	75.13%	100.00%

(19) a. Jij hoeft u daarover niet te schamen of wat *ook*. (WR-P-E-G-0000005399)
'You don't have to be ashamed of that or whatever.'
b. Een fusie met wie *dan ook* is geen optie. (WR-P-P-G-0000599808)
'A fusion with anyone (lit. whoever) is not an option.'

Although *dan ook* is clearly the more frequent option in both corpora, instances with *ook* still account for a considerable proportion of the total. They mainly occur in a specific context, however, namely with indefinite pronouns in comparative constructions. 673 out of 975 occurrences of *ook* in secondary constructions are comparatives (69.03%).

In all 856 comparatives in *SoNaR*, we find a tendency to use the single particle *ook* (n = 673 or 78.62%) rather than the particle combination *dan ook* (n = 183 or 21.38%):

(20) ik weet meer dan *wie ook* over armoede (WR-P-E-A-0000410476)
'I know more than anyone (lit. whoever) about poverty'

According to Hoeksema (2012: 96), the reason for this tendency is that speakers want to avoid a "double *dan*" (i.e. *horror aequi*). Since the comparative particle in Dutch happens to be *dan* (Reinarz/de Vos/de Hoop 2016), speakers tend to prefer *dan wie ook* over *dan wie dan ook*. Moreover, comparative constructions tend to be used with animate pronouns (e.g. *wie* 'who' rather than *wat* 'what'; Hoeksema 2012: 98), and this could explain why the proportion of *ook* is significantly higher in secondary irrelevance constructions with *wie*, while *dan ook* shows a significant preference for inanimate *wat* (Yates $\chi^2 = 601.88$; df = 1; $p < 0.0001$; Cramér's V = 0.39).

Table 13: Chi-square test comparing occurrences of Dutch irrelevance particle(s) (*dan*) *ook* secondary constructions in the *SoNaR* corpus. Standardized residuals are given in brackets, no cells have an expected value below 5

–	ook	dan ook	total
wie	704 **(+16.61)**	822 **(–9.56)**	1,526
wat	271 **(–13.27)**	2,121 **(+7.64)**	2,392
total	975	2,943	3,918

3.2.2 German

The distribution of German irrelevance particles in secondary constructions in the *ConverGENTiecorpus* is shown in Table 14; Table 15 represents their distribution in the *DeReKo*. Examples from the corpora are given in (21) resp. (22).

Table 14: Distribution of irrelevance particles in German secondary constructions in the *ConverGENTiecorpus*

–	*immer*	*auch immer*	*auch*	total
#	6	26	3	35
%	17.14%	74.29%	8.57%	100.00%

(21) a. Weitergabe des Mietgegenstandes an natürliche oder juristische Personen in *welcher* Form *immer* ist dem Mieter untersagt. (GE_Ins_Con_0104)
'A transfer of the rental property to natural or legal persons in whichever form is prohibited to the tenant.'

b. Feigheit, Faulheit, *was auch immer*. (GE_Jou_Com_1040)
'Cowardice, laziness, whatever.'

c. Sollte der Mieter, aus *welchen* Gründen *auch*, seinen Mietvertrag annullieren, erklärt er sich bereit, dem Vermieter Schadenersatz zu erstatten. (GE_Ins_Con_0027)
'Should the tenant, for whichever reasons, cancel their rental contract, they agree to pay the landlord a compensation.'

Table 15: Distribution of irrelevance particles in German secondary constructions in the *DeReKo*

-	immer	immer auch	auch immer	auch	total
#	399	18	4,485	24	4,926
%	8.10%	0.37%	91.05%	0.49%	100.00%

(22) a. Zum Einstieg, zum Verführen, als kleine Zwischenmahlzeit, als *was immer*: Tapas müssen auf den Tisch. (Nürnberger Nachrichten, 8-3-1999)
'As a starter, as a temptation, as a small snack in between, as whatever: there have to be tapas on the table.'

b. Aber wer könnte ein Interesse daran haben, Ihnen *was immer auch* zuzufügen? (Emme, Pierre: Florentinerpakt, 25-3-2011)
'But who could benefit from inflicting anything (lit. whatever) upon you?'

c. Ich bin wichtig. Ich bin... *was auch immer*. (Braunschweiger Zeitung, 23-10-2010)
'I am important. I am ... whatever.'

d. Ob Baldi, Plüss oder *wer auch* sonst: Bern braucht vor allem eines: Den Mut, mit der Vergangenheit zu brechen. (Zürcher Tagesanzeiger, 30-11-1998)
'Whether Baldi, Plüss or whoever else: Bern needs one thing above all: the courage to leave the past behind.'

Whereas *immer* is the most frequent particle in primary irrelevance constructions, it plays only a minor role in secondary constructions. Instead, the latter are clearly dominated by *auch immer*, the only particle (or particle combination) in the *DeReKo* sample that prefers secondary over primary constructions (4,485 secondary constructions out of 6,889 total instances = 65.10%). By contrast, all the other particles and particle combinations clearly prefer primary constructions (*immer*: 399 instances are secondary constructions out of 13,812 total instances = 2.89%; *immer auch*: 18/1,827 = 0.99%; *auch*: 24/772 = 3.11%). This holds especially for the other particle combination, *immer auch*, which does not occur in secondary constructions in the *ConverGENTiecorpus* at all.

4 Differences and similarities

Now that the distributional patterns of irrelevance particles in different construction types have been described in section 3, the most striking differences and similarities between certain particles or particle combinations will be discussed below.

4.1 German *auch* and Dutch *ook*

As mentioned above, the etymologically identical irrelevance particles *auch* and *ook* share their overwhelming rightward tendency. In fact, the distributional tendencies of *auch* and *ook* are strikingly similar (*auch* occupies field IV in 647 out of 669 *W II S IV V*-instances in the *DeReKo* = 96.71%; 4,808/4,977 in the *SoNaR* corpus = 96.60%). Their distributional patterns are statistically identical both in the language-specific corpora and in the subcomponents of the *ConverGENTiecorpus* ($\chi^2 < 0.001$, n = 5,646, df = 1, $p > 0.99$ for the *DeReKo* and *SoNaR* corpus; Fisher's Exact Test: $p > 0.99$ for the *ConverGENTiecorpus*). This rightward tendency has been explained in terms of disambiguation: according to Leuschner (2000: 354), *auch* is more likely to be misinterpreted as a narrow-scope focus particle in field II and more likely to be read as a wide-scope irrelevance particle in field IV (cf. also

Bossuyt/De Cuypere/Leuschner 2018: 110). The same explanation applies to the Dutch particle *ook*.

As has been observed before (Leuschner 2000: 350), occupation of field II by *auch* (and *ook*) is not only much rarer, but also more restricted, as *auch* and *ook* can only occur, at least as irrelevance particles, before lexical subjects, not before pronouns:

(23) a. Was *auch* die Abgeordneten des Bundestags entscheiden – das letzte Wort hat immer wieder das Bundesverfassungsgericht.
(Nürnberger Zeitung, 21-12-2012)
'Whatever the delegates of the Bundestag decide – the Federal Constitutional Court always has the last word.'

b. Wie *ook* zijn medewerkers waren in de regering, bij de Europese Commissie of het Europees Parlement, allen bewaren ze goede herinneringen aan hun vroegere 'baas'. (NE_Lit_Non_1208)
'Whoever his fellow workers were in the government, at the European Commission or the European Parliament, all had good memories to their former 'boss'.'

(24) a. *Was *auch* die entscheiden [...]
'Whatever they decide ...'

b. *Wie *ook* zij waren [....]
'Whoever they were ...'

This positional restriction can be explained by the general tendency of German and Dutch lexical subjects to occupy their base position in [Spec, VP] (Lenerz 1993: 118), i.e. the right periphery of the middle-field, occasionally forcing *auch/ook* to occupy field II despite the above-mentioned risk of ambiguity. According to Behaghel's (1909) "Law of Increasing Constituents" and the principle of end-weight, the preference to occur further to the right is especially strong with lengthier lexical subjects. Conversely, German and Dutch pronouns generally prefer to occupy the left periphery of the middle-field, also known as the "Wackernagel position" (Weiß 2018). Since pronouns are typically thematic, expressing discourse-old, given information, they tend to occur before rhematic, i.e. discourse-new information, which is typically expressed through lexical word classes such as NPs (cf. Noel Aziz Hanna 2015: 46). *Auch* thus never precedes pronouns because its positional preferences are perfectly complementary to those of pronouns.

4.2 German *auch immer* and Dutch *dan ook*

The particle combinations *auch immer* and *dan ook* share four notable similarities. The first is that they are the most frequent option in secondary irrelevance constructions in their respective language, as seen in section 3.2. Although *auch immer* may seem to be more dominant in German secondary constructions (26 out of 35 instances in the *ConverGENTiecorpus* = 74.29%; 4,485/4,926 in the *DeReKo* = 91.05%) compared to Dutch *dan ook* (55/93 in the *ConverGENTiecorpus* = 59.14%; 2,946/3,921 in the *SoNaR* corpus = 75.13%), this difference is only significant in the language-specific corpora (two-tailed two-proportions Z-test $p < 0.0001$), not in the *ConverGENTiecorpus* ($p = 0.11$). It can be explained by the fact that *W auch/immer*-pronouns in comparative constructions do not occur in the German sample at all, while being very frequent in the Dutch sample. As mentioned above (cf. section 3.2.1), it is in this exact context that Dutch secondary irrelevance constructions show a tendency to take the single particle *ook* rather than the particle combination *dan ook*.

Consistent with this similarity, both *auch immer* and *dan ook* are specialized for secondary irrelevance constructions: if all instances of *auch immer* and *dan ook* are considered, a clear majority of them turn out to be secondary constructions. *Dan ook* seems to specialize even more for secondary constructions than *auch immer*: all 55 instances of *dan ook* in the *ConverGENTiecorpus* are secondary constructions, compared to only 26 out of 41 instances with *auch immer* (63.41%). A similar pattern is found in the language-specific corpora (2,946/3,096 in the *SoNaR* corpus = 95.16% vs. 4,485/6,889 = 65.10% in the *DeReKo*; two-tailed two-proportions Z-test: $p < 0.0001$).[14] Thus, whereas German *auch immer* occurs both in secondary irrelevance constructions (where it clearly predominates) and primary constructions, Dutch *dan ook* is almost exclusively found in secondary constructions.

The third similarity of *auch immer* and *dan ook* is that these particle combinations are never broken up by any other constituent, i.e. that the components *auch* and *immer* resp. *dan* and *ook* always occur next to each other. Using terminology suggested by Thurmair (1989: 290) for modal particles, *auch immer* and *dan ook* thus qualifiy as "closed" particle combinations. This suggests that these erstwhile particle combinations have been reanalyzed as single complex particles, enabling them to function as "indefiniteness markers" to the *W*-stem (in the terminology of Haspelmath 1997) in secondary irrelevance constructions.

14 The two-proportions Z-test cannot be performed on the data from the *ConverGENTiecorpus* because the difference between the numerator and denominator is < 5 for the Dutch data (55/55 = 100.00%).

The fourth similarity is the statistically significant preference of *auch immer* and *dan ook* for primary constructions of the *W II/IV V*-type over the *W II S IV V*-type (cf. above, sections 3.1.1 and 3.1.2). Given that *W II/IV V*-constructions have no subject field and therefore tend to be shorter than *W II S IV V*-constructions, the preference of *auch immer* and *dan ook* for shorter or elliptically reduced subordinate clauses (as observed by Leuschner 2000: 353 for *auch immer*) is not surprising.

The most notable difference between *auch immer* and *dan ook* is the complementary nature of their positional tendencies in primary irrelevance constructions. In *W II S IV V*-constructions, *auch immer* shows a strong leftward tendency, occupying field II 954 out of 1,108 total instances in the *DeReKo* (86.10%). *Dan ook*, on the other hand, shows a strong rightward tendency in this construction type (132/136 in field IV in the *SoNaR* corpus = 97.06%). This might seem like a problem, as it has been argued that *auch immer*'s leftward tendency is one of the major factors that caused this particle combination to specialize for secondary constructions (Bossuyt 2016: 64): a *W*-word and one or more subsequent irrelevance particles are more likely to be reanalyzed as a new unit if the particles typically occur in immediate adjacency to the *W*-word. This factor may well apply to German *auch immer*, but it is obviously irrelevant for *dan ook*, given that Dutch does not have any irrelevance particles with a leftward tendency to begin with. The reason that *dan ook* specializes for secondary constructions rather than *ook* alone is due to the fact that a complex particle is less prone to ambiguity as an indefiniteness marker than a single particle. This is especially true in clause-medial contexts, in which secondary constructions often occur. For the very same reason, the complex particle *auch immer* is more frequent in German secondary constructions than *immer*, which also has a preference for field II, but is a single particle instead of a closed particle combination (ibid.).

4.3 German *immer* and English *-ever*

Not only is German *immer* related to English *-ever* etymologically to the extent that the initial *i-* in *immer* is cognate with the *e-* in English *ever* (Leuschner 1996), its leftward tendency is reminiscent of the positional shift undergone by *ever* in the history of English. In present-day English, attaching itself to the *WH*-word is the only option for *-ever* (cf. above). For *immer*, it is almost the only option: *immer* occupies field II in 6,075 out of 6,114 *W II S IV V*-instances in the *DeReKo* sample (99.36%). Although *immer* competes for this position with pronominal subjects (cf. section 4.1 above), the pronoun has successfully forced *immer* to occupy field IV in only 39 instances in the entire *DeReKo* sample. Since only pronouns com-

pete for Wackernagel's position, *immer* never occurs behind lexical subjects, as shown by (26) in comparison with the original in (25):

(25) Und was es *immer* gewesen sein mag: Der Verdächtige ist nicht vorbestraft und erst recht nicht verurteilt. (Nürnberger Nachrichten, 16-5-1998)
'And whatever it may have been: the suspect has not been previously convicted and surely never been sentenced.'

(26) * Was das Verbrechen *immer* gewesen sein mag [...]
'Whatever the crime may have been ...'

As shown by Leuschner (2001; 2006: 134–140), *immer* and *ever* first occurred in irrelevance constructions as free-choice adverbs supporting the quantificational effect of the semantically opaque irrelevance markers *so ... so* (e.g. Old English *swa hwylc swa* 'whoever, whichever'; Old High German *so wér so* 'whoever'). *Immer* and *ever* then began replacing *so ... so* as the main irrelevance marking strategy, a grammaticalization process which was accompanied by the omission of the left-hand *swa* (> *so*) in English, eventually resulting in *WH*-(*so*)-*ever*-compounds, and the right-hand *so* (> *s*-) in German, resulting in combinations like *swâ iemer* 'wherever'. English *-so-* was eventually left out completely in irrelevance constructions (for *whatsoever*, which still occurs as a post-nominal intensifier, cf. above and Leuschner 2001), and German irrelevance marking *s-W*-words collapsed with bare *W*-words in the 14th century (Leuschner 2006: 135), leaving *iemer* (> *immer*) and *auch* as clause-internal irrelevance marking. Both *ever* and *immer* occurred initially in field IV, i.e. in the typical position of adverb, but following their reanalysis as quantificational particles began shifting towards field II as *so ... so* became increasingly obsolete and the new strategies of irrelevance marking became more and more obligatory (cf. Leuschner 2006 and Bossuyt/De Cuypere/Leuschner 2018 for more details).

While *immer* and *-ever* both underwent grammaticalization, this process happened much faster in English than in German. The last instances with *æure* (> *ever*) in field IV seem to be attested around the 12th century:

(27) Luue ðine nexte al swa ðe seluen, hwat manne swa he *æure* bie! (cited in Leuschner 2006: 135)
'Love thy neighbour like thyself, whatever man he be!'

In German, however, the positional tendencies of *immer* and *auch* did not emerge clearly until well into the 19th century (Leuschner 2006: 136), as suggested by verses like (28) from Johann Wolfgang von Goethe (1749–1832):

(28) Und man kommt in's Gered', wie man sich *immer* stellt. (cited in Goethe's *Faust I*, line 3201)
'And one becomes the subject of gossip, however one (lit.: how one ever) positions oneself.'

Unlike with *-ever*, the grammaticalization of *immer* is still incomplete. While phrases like *whichever house* and *however beautiful* are perfectly grammatical in English, their German equivalents with *immer* are ungrammatical or at least highly unusual: *welches* *(*immer*) *Haus* ?(*immer*), *wie* *(*immer*) *schön* ?(*immer*). When *wessen* 'whose' modifies an intervening NP, as in (29a), *immer* is ruled out, but *auch immer* is allowed.[15] When *wessen* functions as a genitive object, by contrast, and no constituents intervene between the *W*-word and the particle as in (29b), *immer* is unproblematic:

(29) a. mit wessen Geld *auch immer* [**immer*] sie bezahlt wurden
(St. Galler Tagblatt, 18-3-2010)
'with whoever's money they got payed'
b. wessen *immer* man mich anklagt (Süddeutsche Zeitung, 31-3-1998)
'Whatever (some)one accuses me of'

Immer also seems to be problematic with complex *W*-words such as *womit* 'with what/which' (lit. 'where-with'), as suggested by Leuschner (2000: 350). These restrictions have so far prevented *immer* from becoming the sole irrelevance particle in German and attaining univerbation with the *W*-word, as has happened in English. Its obligatorification seems to be counteracted by the presence of other particles, as the above-mentioned restrictions are more likely to encourage the use of *auch* or particle combinations rather than *immer* alone in these specific contexts.

5 Conclusion and prospects

The present study has documented and analyzed the distributional patterns of the irrelevance particles *-ever*, *immer* and/or *auch* and (*dan*) *ook* in both primary and secondary irrelevance constructions. A contrastive corpus triangulating approach

15 (29a) would be grammatical with *auch immer* in either field II or field IV, or, alternatively, with *auch* in field IV. In any case, *immer* in field II is ruled out.

was adopted, thereby expanding the scope of a previous study by Bossuyt/De Cuypere/Leuschner (2018) and providing a semi-replication of Leuschner (2000).

From a diachronic perspective, the synchronic analysis can be read as a snapshot of a long-term process of emergence-by-grammaticalization. As far as primary constructions are concerned, this is nearly completed in English, where *-ever* is the sole irrelevance particle and only occurs in univerbation with the *WH*-word. There are only a few small defects in the *WH-ever*-paradigm, like the **whyever* gap (Leuschner 2006: 41) and residual *-so-* in intensifying *whatsoever*. In German, the grammaticalization process is not only incomplete, but seems to have lost its former directionality: although *immer* shows a very similar leftward tendency to *-ever*, it has not yet reached univerbation with the *W*-word and seems unlikely to do so in the foreseeable future because its obligatorification is hindered by the systemic presence of both *auch* and the particle combination *auch immer*. Thus, the *W immer/auch*-paradigm seems to be stuck in an uneasy balance. Dutch *W … ook* shows only weak signs of grammaticalization: although *ook* did undergo function-specific semantic changes when it was recruited from the focus particle, and shows a clear preference for field IV (cf. Bossuyt 2016: 59 and Leuschner 2013: 53 on German *auch*), its position in this field is not absolute and the result of its preferential position is precisely to make it discontinuous with the *W*-word. *Ook* thus fails to show even the most rudimentary signs of coalescence (Lehmann 2015: 157–167), a clear indication that any further increase in grammaticalization is blocked.

In secondary constructions, however, we see a different pattern. Dutch *W dan ook* is highly specialized for secondary constructions and the most functionally versatile of all three languages, occurring as a discourse marker, general extender, and indefinite pronoun. German *W auch immer* occurs frequently in the first two functions, but is still rare as an indefinite pronoun. The same is true for English *WH-ever* in secondary constructions, mainly due to the systemic presence of the *any*-series.

The subsystem of irrelevance marking through particles thus participates in the larger "grammaticalization building-site" of concessive conditionality in English, German, and Dutch (Leuschner 2006). Follow-up research should look at the interaction, both in terms of quantification (i.e. semantics) and of surface distribution, between irrelevance particles on the one hand and expressions of epistemic modality, particularly *may/mögen/mogen*, and the present subjunctive as strategies of free-choice quantification on the other hand:

(30) a. It might have something to do with people trying to express their frustration -- what*ever* that *may* be.
(COCA, NEWS: Atlanta)

b. Steht am Schluss eines Artikels "pd", hat nicht die Zeitungsredaktion geschrieben, sondern der "Pressedienst", wer *immer* das sein *möge*. (St. Galler Tagblatt, 11-5-2012)
'If it says "pd" at the end of an article, then the newspaper editorial didn't write [it], but the "press-service", whoever that may be.'
c. Wat Henin *ook moge* beweren, zij start als favoriete. (WR-P-P-G-0000237815)
'Whatever Henin may claim, she starts as the favorite.'

Another avenue is to investigate the alteration between clause-internal strategies of irrelevance marking, i.e. through irrelevance particles, and clause-external strategies such as elliptical expressions of irrelevance, which come to the building-site with a grammaticalization history of their own (Leuschner 2006). As Bossuyt/De Cuypere/Leuschner (2018: 117) demonstrate, rare instances of overlap between clause-external and clause-internal strategies exist in German, and equivalent examples occur in Dutch:

(31) a. *Egal*, was sie *auch* tun (die tageszeitung, 2-12-2006)
'No matter what (lit. whatever) they do'
b. Het is het nie[t] waard jong, *gelijk* met wie je *ook* zo een one-night-stand zou willen doen. (WR-P-E-A-0000047811)
'It's not worth it, man, no matter with whom (lit. whomever) you would like to have a one night stand.'

The language-specific and cross-linguistic patterning of such an overlap remains to be seen. Equivalent structures in English would feature a *WH-ever*-word in combination with *no matter* (or some other clause-external marker). The fact that there are no such examples in the BYU-sample at all matches the observation that the overlap occurs in German mainly with the rightward-tending *auch*, but only rarely with the leftward-tending *immer*. Any future investigation taking into account irrelevance markers other than clause-internal particles is thus likely to confirm the position of German irrelevance marking strategies in between those of Dutch and English.

Corpora

Davies, Mark: BYU-BNC. (Based on the British National Corpus from Oxford University Press). Available online at https://corpus.byu.edu/bnc/.
Davies, Mark: The Corpus of Contemporary American English (COCA): 560 million words, 1990-present. Available online at https://corpus.byu.edu/coca/.

Davies, Mark: The Strathy Corpus: 50 million words of Canadian English. Available online at https://corpus.byu.edu/can.
Davies, Mark: Hansard Corpus: part of the SAMUELS project. Available online at www.hansard-corpus.org/.
Davies, Mark: The Wikipedia Corpus: 4.6 million articles, 1.9 billion words. Adapted from Wikipedia. Available online at https://corpus.byu.edu/wiki/.
Instituut voor de Nederlandse taal: OpenSoNaR. Available online at http://opensonar.inl.nl/.
Leibniz-Institut für Deutsche Sprache, Mannheim: Das Deutsche Referenzkorpus DeReKo. Available online at www.ids-mannheim.de/kl/projekte/korpora/.
Lauwers, Peter/Plevoets, Koen: ConverGENTiecorpus. Institutionally available at Ghent University.

References

Behaghel, Otto (1909): Beziehungen zwischen Umfang und Reihenfolge von Satzgliedern. In: Indogermanische Forschungen 25. 110–142.
Bossuyt, Tom (2016): Zur Distribution von Irrelevanzpartikeln in *was immer*/*auch*-Konstruktionen: Positionelle und kombinatorische Varianz im Deutschen Referenzkorpus. In: Germanistische Mitteilungen 42, 1. 45–70.
Bossuyt, Tom/De Cuypere, Ludovic/Leuschner, Torsten (2018): Emergence phenomena in German *W-immer*/*auch*-subordinators. In: Fuß, Eric/Konopka, Marek/Trawiński, Beata/Waßner, Ulrich H. (eds.): Grammar and corpora 2016. Heidelberg: Heidelberg University Publishing. 97–120.
Breindl, Eva (2014): Irrelevanzkonditionale Konnektoren. In: Breindl, Eva/Volodina, Anna/Waßner, Ulrich H. (eds.): Handbuch der deutschen Konnektoren. Vol. 2: Semantik der deutschen Satzverknüpfer. (= Schriften des Instituts für Deutsche Sprache 13). Berlin/New York: De Gruyter. 964–1009.
Brinton, Laurel J. (2017): The evolution of pragmatic markers in English: Pathways of change. Cambridge: Cambridge University Press.
Haspelmath, Martin (1997): Indefinite pronouns. Oxford: Oxford University Press.
Haspelmath, Martin/König, Ekkehard (1998): Concessive conditionals in the languages of Europe. In: van der Auwera, Johan (ed.): Adverbial constructions in the languages of Europe. (= Empirical Approaches to Language Typology 20). Berlin/New York: De Gruyter. 563–641.
Hentschel, Elke (2012): Abtönungspartikeln: Die Läuse im Pelz der Sprache. In: Leupold, Gabriele/Passet, Eveline (eds.): Im Bergwerk der Sprache: Eine Geschichte des Deutschen in Episoden. Göttingen: Wallstein. 124–142.
Hoeksema, Jacob (2012): *Wie dan ook, wat dan ook,* etc. als free choice indefinites en negatiefpolaire uitdrukkingen. In: TABU. Bulletin voor Taalwetenschap 40, 3/4. 89–109.
Kadmon, Nirit/Landman, Fred (1993): *Any*. In: Linguistics and Philosophy 16. 353–422.
König, Ekkehard (1986): Conditionals, concessive conditionals and concessives: Areas of contrast, overlap and neutralization. In: Traugott, Elizabeth C./Ter Meulen, Alice/Snitzer Reilly, Judy/Ferguson, Charles A. (eds.): On conditionals. Cambridge: Cambridge University Press. 229–246.
König, Ekkehard (2010): Dimensionen der Bedeutung und Verwendung von Modalpartikeln im Deutschen: Grundlagen einer Bestandsaufnahme. In: Harden, Theo/Hentschel, Elke (eds.): 40 Jahre Partikelforschung. (= Stauffenburg Linguistik 55). Tübingen: Stauffenburg. 79–96.

König, Ekkehard/Eisenberg, Peter (1984): Zur Pragmatik von Konzessivsätzen. In: Stickel, Gerhard (ed.): Pragmatik in der Grammatik: Jahrbuch 1983 des Instituts für deutsche Sprache. (= Sprache der Gegenwart 60). Düsseldorf: Schwann. 313–332.

Kupietz, Marc/Lüngen, Harald (2014): Recent developments in DeReKo. In: Calzolari, Nicoletta/Choukri, Khalid/Declerck, Thierry/Loftsson, Hrafn/ Maegaard, Bente/Mariani, Joseph/Moreno, Asuncion/Odijk, Jan/Piperidis, Stelios (eds.): Proceedings of the Ninth International Conference on Language Resources and Evaluation: LREC 2014. Reykjavík: European Language Resources Association. 2378–2385.

Lehmann, Christian (1984): Der Relativsatz: Typologie seiner Strukturen, Theorie seiner Funktionen, Kompendium seiner Grammatik. (= Language Universals Series 3). Tübingen: Narr.

Lehmann, Christian (2015): Thoughts on grammaticalization. 3rd edition. (= Classics in Linguistics 1). Berlin: Language Science Press.

Lenerz, Jürgen (1993): Zu Syntax und Semantik deutscher Personalpronomina. In: Reis, Marga (ed.): Wortstellung und Informationsstruktur. (= Linguistische Arbeiten 306). Tübingen: Niemeyer. 117–154.

Leuschner, Torsten (1996): Ever and universal quantifiers of time: Observations from some Germanic languages. In: Language Sciences 18. 469–484.

Leuschner, Torsten (2000): '..., *wo immer* es mir begegnet, ... – *wo* es *auch* sei': Zur Distribution von 'Irrelevanzpartikeln' in Nebensätzen mit *w- auch*/*immer*. In: Deutsche Sprache 28. 342–356.

Leuschner, Torsten (2001): Nebensatzkonnektoren des Typs 'W-Wort + Partikel(n)' (Deutsch wer auch immer usw.) im Germanischen: Eine intragenetische Typologie aus areallinguistischer Sicht. In: Studia Germanica Gandensia 2. 3–26.

Leuschner, Torsten (2005): Nonspecific free relatives and (anti)grammaticalization in English and German. In: Folia Linguistica Historica 26, 1-2. 45–69.

Leuschner, Torsten (2006): Hypotaxis as building-site: The emergence and grammaticalization of concessive conditionals in English, German and Dutch. Munich: LINCOM Europa.

Leuschner, Torsten (2013): Was Partikeln wohl (auch immer) mit Indifferenz zu tun haben: Funktionale und linguistikdidaktische Perspektiven. In: Germanistische Mitteilungen 39, 1. 37–62.

Müller, Sonja (2017): Alte und neue Fragen der Modalpartikel-Forschung. In: Linguistische Berichte 252. 383–441.

Noel Aziz Hanna, Patrizia (2015): Wackernagels Gesetz im Deutschen: Zur Interaktion von Syntax, Phonologie und Informationsstruktur. (= Studia Linguistica Germanica 122). Berlin/Boston: De Gruyter.

Nübling, Damaris (2005): Von *in die* über *in'n* und *ins* bis *im*: Die Klitisierung von Präposition und Artikel als 'Grammatikalisierungsbaustelle'. In: Leuschner, Torsten/Mortelmans, Tanja/De Groodt, Sarah (eds.): Grammatikalisierung im Deutschen. (= Linguistik – Impulse & Tendenzen 9). Berlin: De Gruyter. 105–131.

Oostdijk, Nelleke/Reynaert, Martin/Hoste, Véronique/Schuurman, Ineke (2013): The construction of a 500-million-word reference corpus of contemporary written Dutch. In: Spyns, Peter/Odijk, Jan (eds.): Essential speech and language technology for Dutch: Results by the STEVIN-programme. Berlin/Heidelberg: Springer. 219–247.

Overstreet, Maryann (1999): Whales, candlelight, and stuff like that: General extenders in English discourse. Oxford/New York: Oxford University Press.

Reinarz, Lukas/de Vos, Hugo/de Hoop, Helen (2016): Conflicting constraints in the comparative cycle. In: Journal of Germanic Linguistics 28, 4. 403–425.

Reiners, Ludwig (1949): Stilkunst: Ein Lehrbuch deutscher Prosa. Munich: Biederstein.
Thurmair, Maria (1989): Modalpartikeln und ihre Kombinationen. (= Linguistische Arbeiten 223). Tübingen: Niemeyer.
van Haeringen, Coenraad B. (1956): Nederlands tussen Duits en Engels. The Hague: Servire.
Waßner, Ulrich (2006): Zur Relevanz von und zur Irrelevanz bei Irrelevanzkonditionalen. In: Breindl, Eva/Gunkel, Lutz/Strecker, Bruno (eds.): Grammatische Untersuchungen: Analysen und Reflexionen. Gisela Zifonun zum 60. Geburtstag. (= Studien zur Deutschen Sprache 36). Tübingen: Narr. 381–399.
Weiß, Helmut (2018): The Wackernagel complex and pronoun raising. In: Agnes Jäger/Gisella Ferraresi/Helmut Weiß (eds.): Clause structure and word order in the history of German. Oxford: Oxford University Press. 132–154.
Wöllstein, Angelika (2014): Topologisches Satzmodell. 2nd updated edition. Heidelberg: Winter.